JOURNAL OF RESEARCH of the National Bureau of Standards — B. Mathematical Sciences
Vol. 74B, No. 1, January–March 1970

Bounds on a Polynomial*

T. J. Rivlin**

(December 15, 1969)

Methods for computing the maximum and minimum of a polynomial with real coefficients in the interval [0, 1] are described, and certain bounds are given.

Key words: Bernstein polynomials; bounds; polynomials.

Introduction

Let $p(x)=a_0+a_1x+\ldots+a_nx^n$ have real coefficients. If I is the interval [0, 1] and

$$M=\max_{x\epsilon I} p(x), \qquad m=\min_{x\epsilon I} p(x),$$

we wish to study some methods of approximating m and M relatively easily. Apart from their intrinsic interest, such methods would seem to have application in computations using interval arithmetic in which a basic operation is the determination of the range of a rational function with a given interval as domain. The restriction of our discussion to the interval I entails no loss of generality, since any interval $[a, b]$ can be mapped onto I by a linear function of x, and such changes of variable leave the set of polynomials of degree at most n invariant.

Some estimates of m and M were given by Cargo and Shisha.[1] They first observe that if $n \geqslant 1$ and $J=\{1, \ldots, k\}$,

$$\min_{j\epsilon J} p\left(\frac{j}{k}\right)-\frac{1}{k}\sum_{j=1}^{n} j|a_j| \leqslant m \leqslant M \leqslant \max_{j\epsilon J} p\left(\frac{j}{k}\right)+\frac{1}{k}\sum_{j=1}^{n} j|a_j|. \tag{1}$$

Then they note that p has a representation in Bernstein form, namely,

$$p(x)=\sum_{j=0}^{n} b_j \binom{n}{j} x^j(1-x)^{n-j}, \tag{2}$$

which leads immediately to the bounds

$$\min_{j\epsilon N} b_j \leqslant m \leqslant M \leqslant \max_{j\epsilon N} b_j, \tag{3}$$

where $N=\{0, \ldots, n\}$.

*An invited paper.
**Present address: IBM Research Center, Yorktown Heights, New York 10598.
[1] Cargo, G. T., and Shisha, O., The Bernstein form of a polynomial, J. Res. Nat. Bur. Stand. (U.S.), **70B** (Math. Sci.) No. 1, 79–81 (Jan–Mar. 1966).

The numbers $b_0, \ldots, b_n$ are determined by

$$b_j = \sum_{i=0}^{j} a_i \frac{\binom{j}{i}}{\binom{n}{i}}, \qquad j=0, \ldots, n. \tag{4}$$

Our work is intended to be a sequel to the paper of Cargo and Shisha (see footnote 1). In section 1, we give a sharper result than (1). Section 2 is devoted to a generalization of the representation (2) and a subsequent improvement of (3), including estimates of the precision of the bounds obtained for m and M, which are due, essentially, to S. Bernstein.[2] The paper concludes with section 3, containing some results in the case that the coefficients $a_0, \ldots, a_n$ are complex numbers.

1. Estimates Using Function Values

THEOREM 1: *Suppose* $0=t_0<t_1< \ldots <t_k=1$, *and* $d_k=\max(t_{j+1}-t_j)$, $j=0, \ldots, k-1$, *then, if* $K=\{0, \ldots, k\}$

$$\min_{i\in K} p(t_i) - \frac{d_k^2}{8}\max_{x\in I}|p''(x)| \leq m \leq M \leq \max_{i\in K} p(t_i) + \frac{d_k^2}{8}\max_{x\in I}|p''(x)|. \tag{5}$$

PROOF: Suppose $p(\xi)=M$ and

$$|t_j-\xi| \leq |t_i-\xi|,\ i=0, \ldots, k,$$

so that

$$|t_j-\xi| \leq \frac{d_k}{2}\,.$$

According to Taylor's formula, we have

$$p(t_j)=p(\xi)+(t_j-\xi)p'(\xi)+\frac{(t_j-\xi)^2}{2}p''(\eta) \tag{6}$$

where $\eta\epsilon I$. If $\xi=0, 1$, the right-most inequality in (5) is trivially true. If $0<\xi<1$ then $p'(\xi)=0$ and (6) implies that

$$M \leq \max_{j\epsilon K} p(t_j) + \frac{(t_j-\xi)^2}{2}\max_{\eta\epsilon I}|p''(\eta)| \leq \max_{j\epsilon K} p(t_j)+\frac{d_k^2}{8}\max_{\eta\epsilon I}|p''(\eta)|.$$

An entirely analogous argument establishes the lower bound on m in (5).

COROLLARY: *If* $t_i=i/k$ *then* $d_k=1/k$, *and since*

$$\max_{x\epsilon I}|p''(x)| \leq \sum_{j=0}^{n}(j-1)j|a_j|$$

(5) *now becomes,*

$$\min_{j\epsilon K} p\left(\frac{j}{k}\right)-\frac{1}{8k^2}\sum_{j=0}^{n}(j-1)j|a_j| \leq m \leq M \leq \max_{j\epsilon K} p\left(\frac{j}{k}\right)+\frac{1}{8k^2}\sum_{j=1}^{n}(j-1)j|a_j|. \tag{7}$$

[2] Bernstein, S. N., Collected Works, Vol. I, Translation: AEC-tr-3460, pp. 68–73.

The improvement of the error estimate in (7) over that in (1) is evident.

If we consider $p(x)=x-x^2$ then $M=1/4$, which is precisely the upper bound given by (7) with $k=1$.

2. The Bernstein Form

Suppose $s=0, 1, \ldots, k$ then we have

LEMMA 1:

$$x^s = \sum_{j=s}^{k} \frac{\binom{j}{s}}{\binom{k}{s}} \binom{k}{j} x^j (1-x)^{k-j}.$$

PROOF:

$$x^s = x^s(x+(1-x))^{k-s} = \sum_{i=0}^{k-s} \binom{k-s}{i} x^{s+i}(1-x)^{k-s-i}$$

$$= \sum_{j=s}^{k} \binom{k-s}{j-s} x^j(1-x)^{k-j} = \sum_{j=s}^{k} \frac{\binom{j}{s}}{\binom{k}{s}} \binom{k}{j} x^j(1-x)^{k-j}.$$

Thus, the functions

$$\binom{k}{j} x^j(1-x)^{k-j},\ j=0, \ldots, k$$

span the space, P_k, of polynomials of degree at most k and are linearly independent. Hence, if $p(x)=a_0+\ldots+a_n x^n$ (a_j are arbitrary numbers) and $k \geqslant n$ then

$$p(x) = \sum_{j=0}^{k} b_j^{(k)} \binom{k}{j} x^j(1-x)^{k-j} \tag{8}$$

where

$$b_j^{(k)} = \sum_{s=0}^{j} a_s \frac{\binom{j}{s}}{\binom{k}{s}},\ j=0, \ldots, k, \tag{9}$$

with the assumption that $a_s=0$ for $s>n$.

We call (8) the generalized Bernstein form of p.

THEOREM 2: *If* $p(x)=a_0+\ldots+a_n x^n$ *has real coefficients then for each* $k \geqslant n$ *we have (putting* $K=\{0, \ldots, k\}$)

$$b(k) = \min_{j\in K} b_j^{(k)} \leqslant m \leqslant M \leqslant \max_{j\in K} b_j^{(k)} = B(k).$$

PROOF: In view of (8) for each $x \epsilon I$, $p(x)$ is a convex combination of $b_0^{(k)}, \ldots, b_k^{(k)}$.

Following S. Bernstein, we can bound the discrepancies $B(k)-M$ and $m-b(k)$ as follows. We recall that if $f(x)$ is defined on I

$$B_k(f; x) = \sum_{j=0}^{k} f\left(\frac{j}{k}\right) \binom{k}{j} x^j(1-x)^{k-j}.$$

Suppose $0 \leqslant s \leqslant n$. Then $B_k(x^s; x) - x^s \epsilon P_k$ and hence

$$B_k(x^s; x) - x^s = \sum_{j=0}^{k} \delta_j(s) \binom{k}{j} x^j (1-x)^{k-j}.$$

THEOREM 3: *If* $k \geqslant n \geqslant 1$

$$\delta_j(s) \leqslant \frac{(s-1)^2}{k}; \; j=0, \ldots, k; \; s=0, \ldots, n. \tag{10}$$

PROOF. Since $B_k(1; x) \equiv 1$ and $B_k(x; x) \equiv x$ we have

$$\delta_j(0) = \delta_j(1) = 0, \; j=0, \ldots, k.$$

Henceforth, we assume $s \geqslant 2$.

(i) $0 \leqslant j < s$.

In view of Lemma 1, we have

$$\delta_j(s) = \left(\frac{j}{k}\right)^s \leqslant \left(\frac{s-1}{k}\right)^s \leqslant \left(\frac{s-1}{k}\right)^2 \leqslant \frac{(s-1)^2}{k}.$$

(ii) $2 \leqslant s \leqslant j$

$$\delta_j(s) = \left(\frac{j}{k}\right)^s - \frac{j!(k-s)!}{(j-s)!k!} = \left(\frac{j}{k}\right)^s - \frac{j(j-1) \ldots (j-(s-1))}{k(k-1) \ldots (k-(s-1))}$$

$$= \left(\frac{j}{k}\right)^s \left[1 - \frac{\left(1-\frac{1}{j}\right) \ldots \left(1-\frac{s-1}{j}\right)}{\left(1-\frac{1}{k}\right) \ldots \left(1-\frac{s-1}{k}\right)}\right]$$

$$\leqslant \left(\frac{j}{k}\right)^s \left[1 - \left(1-\frac{1}{j}\right) \ldots \left(1-\frac{s-1}{j}\right)\right]$$

$$\leqslant \left(\frac{j}{k}\right)^s \left[1 - \left(1-\frac{s-1}{j}\right)^{s-1}\right].$$

Applying the mean value theorem to $(1-x)^{s-1}$, we obtain

$$1 - \left(1-\frac{s-1}{j}\right)^{s-1} \leqslant \frac{(s-1)^2}{j},$$

hence

$$\delta_j(s) \leqslant \frac{(s-1)^2}{k} \left(\frac{j}{k}\right)^{s-1} \leqslant \frac{(s-1)^2}{k}. \tag{11}$$

REMARK: If $k \geqslant 2$, (10) can be improved slightly to

$$\delta_j(s) \leqslant \left(1-\frac{1}{k}\right) \frac{(s-1)^2}{k}, \tag{12}$$

since:

(i) If $j < s$,

$$\delta_j(s) \leqslant \frac{(s-1)^2}{k} \cdot \frac{1}{k} \leqslant \left(1-\frac{1}{k}\right) \frac{(s-1)^2}{k}.$$

(ii) If $s \leqslant j \leqslant k$,

$$\delta_k(s) = 0,$$

and so (11) can be restricted to $j < k$ leading to

$$\delta_j(s) \leqslant \frac{(s-1)^2}{k} \cdot \left(\frac{k-1}{k}\right)^s \leqslant \frac{(s-1)^2}{k} \cdot \left(\frac{k-1}{k}\right).$$

THEOREM 4: *If* $p(x) = a_0 + \ldots + a_n x^n$, *and* $k \geqslant n \geqslant 1$ *then for* $j = 0, \ldots, k$

$$p\left(\frac{j}{k}\right) - b_j^{(k)} = \delta_j \tag{13}$$

satisfies

$$|\delta_j| \leqslant \frac{A}{k} \tag{14}$$

where

$$A = \sum_{s=2}^{n} (s-1)^2 \, |a_s|. \tag{15}$$

If $k \geqslant 2$, *then*

$$|\delta_j| \leqslant A \frac{k-1}{k^2}, \; j = 0, \ldots, k.$$

PROOF: Since $B_k(p; x) - p(x) = \sum_{j=0}^{k} \delta_j \binom{k}{j} x^j (1-x)^{k-j}$, while in view of Theorem 3

$$B_k(p; x) - p(x) = \sum_{j=0}^{k} \sum_{s=0}^{k} a_s \delta_j(s) \binom{k}{j} x^j (1-x)^{k-j},$$

then

$$\delta_j = \sum_{s=0}^{k} a_s \delta_j(s),$$

and the Theorem follows from (10) and (12).

COROLLARY: *If* $a_0, \ldots, a_n$ are real, then

$$B(k) - M \leqslant \frac{A}{k}; \; m - b(k) \leqslant \frac{A}{k}. \tag{16}$$

372-985 O - 70 - 5

If k $\geqslant 2$

$$B(k)-M \leqslant A\frac{k-1}{k^2};\ m-b(k) \leqslant A\frac{k-1}{k^2}. \tag{17}$$

Thus, $B(k)$ converges to M as $k \to \infty$, and to determine how large we need to choose k in order to be a given distance away from M, we need only consult (16) or (17) (analogously for $b(k)$ and m). Cargo and Shisha (see footnote 1) give a difference table method for calculating $b_1^{(n)}, \ldots, b_{n-1}^{(n)}$ which can be used equally well to find $b_1^{(k)}, \ldots, b_{k-1}^{(k)}$ ($b_0^{(k)} = a_0$; $b_k^{(k)} = a_0 + \ldots + a_n$, always).

Remark: Once a k has been chosen and the numbers $b_j^{(k)}$ ordered according to size, it may be possible to improve on $B(k)$ as an estimate for M, by means of a final correction. Suppose

$$B(k) = b_i^{(k)},\ i \neq 0,\ k,$$

and

$$\max_{j \neq i} b_j^{(k)} = b_l^{(k)} < b_i^{(k)}.$$

Put

$$p_j(x) = \binom{k}{j} x^j (1-x)^{k-j},\ j = 0, \ldots, k$$

then $0 \leqslant p_i(x) < 1$ for all $x \epsilon I$, and

$$p(x) = p_i(x) b_i^{(k)} + (1-p_i(x)) \sum_{\substack{j=0 \\ j \neq i}}^{k} b_j^{(k)} \binom{k}{j} \frac{x^j (1-x)^{k-j}}{1-p_i(x)}. \tag{18}$$

Since

$$\frac{p_j(x)}{1-p_i(x)} \geqslant 0$$

and

$$\sum_{\substack{j=0 \\ j \neq i}}^{k} \frac{p_j(x)}{1-p_i(x)} = 1,$$

we conclude from (18) that

$$p(x) \leqslant p_i(x) b_i^{(k)} + (1-p_i(x)) b_l^{(k)} = (b_i^{(k)} - b_l^{(k)}) p_i(x) + b_l^{(k)}.$$

Since $p_i(x)$ assumes its maximum on I at $x = i/k$, we obtain, finally,

$$M \leqslant p_i\left(\frac{i}{k}\right) b_i^{(k)} + \left(1 - p_i\left(\frac{i}{k}\right)\right) b_l^{(k)} < B(k).$$

An analogous result holds for m.

3. The Complex Case

Suppose the coefficients $a_0, \ldots, a_n$ are complex numbers. Let C be the convex hull of $p(I)=\{p(x) : x\epsilon I\}$, and let C_k be the convex hull of $\{b_0^{(k)}, \ldots, b_k^{(k)}\}$. Then in view of (8), we have, for each $k \geqslant n$

$$C \subseteq C_k.$$

We have

LEMMA 2: *For each* k $\geqslant$ n

$$C_{k+1} \subseteq C_k.$$

PROOF: It is an easy consequence of (9) that $b_0^{(k)}=a_0$, $b_k^{(k)}=a_0+\ldots+a_n$ and

$$b_j^{(k+1)}=\left(1-\frac{j}{k+1}\right)b_j^{(k)}+\frac{j}{k+1}b_{j-1}^{(k)};\ j=1, \ldots, k.$$

Thus

$$b_j^{(k+1)}\epsilon C_k,\ j=0, \ldots, k+1$$

and the lemma follows.

Note that C and C_k, $k=n, n+1, \ldots$ are each compact sets in the plane, and that

$$C \subseteq \bigcap_{k=n}^{\infty} C_k. \tag{19}$$

Indeed,

THEOREM 5:

$$C=\bigcap_{k=n}^{\infty} C_k.$$

PROOF. Suppose there exists

$$z\epsilon \bigcap_{k=n}^{\infty} C_k$$

and $z \notin C$. Since C is compact, dist $(z, C)=d>0$. Choose $k \geqslant n$ so large that $(A/k)<d$. Now, there exist $\lambda_0, \ldots, \lambda_k$ satisfying $\lambda_j \geqslant 0$, $\Sigma\lambda_j=1$ such that

$$z=\sum_{j=0}^{k} \lambda_j b_j^{(k)},$$

hence, in view of (13) and (14)

$$\sum_{j=0}^{k} \lambda_j p\left(\frac{j}{k}\right)-z=\sum_{j=0}^{k} \lambda_j \delta_j,$$

and

$$\left|\sum_{j=0}^{k} \lambda_j p\left(\frac{j}{k}\right)-z\right| \leqslant \frac{A}{k}<d. \tag{20}$$

But

$$\sum_{j=0}^{k} \lambda_j p\left(\frac{j}{k}\right) \epsilon C,$$

and (20) contradicts dist $(z, C) = d$, thus proving the theorem.

Note that the same proof shows that

$$\text{dist } (C_k, C) \leqslant \frac{A}{k}.$$

(Paper 74B1–318)